JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 90, 64–71 (1982)

A Min Max Theorem*

RAUL F. MANASEVICH†

Universidad de Los Andes, Facultad de Ciencias, Departamento de Matematica, Merida, Edo Merida, Venezuela

Submitted by George Leitmann

1. Introduction

The purpose of this paper is to generalize a Max–Min theorem due to Lazer *et al.* [4]. Their theorem is as follows:

THEOREM 1. *Let X and Y be two closed subspaces of a real Hilbert space H such that X is finite dimensional and $H = X \oplus Y$ (X and Y not necessarily orthogonal). Let $f\colon H \to \mathbb{R}$ be a C^2 functional and let ∇f and $D^2 f$ denote the gradient and Hessian of f, respectively. Suppose that there exist two positive constants m_1 and m_2 such that*

$$\langle D^2 f(u)h, h\rangle \leqslant -m_1 \|h\|^2$$

$$\langle D^2 f(u)k, k\rangle \geqslant m_2 \|k\|^2,$$

for all $u \in H$, $h \in X$, and $k \in Y$. Then f has a unique critical point, i.e., there exists a unique $v_0 \in H$, such that, $\nabla f(v_0) = 0$. Moreover, this critical point is characterized by the equality

$$f(v_0) = \max_{x \in X} \min_{y \in Y} f(x + y).$$

In this paper we generalize this result to the case where X and Y are not necessarily finite dimensional. Also our conditions on the Hessian of f are less restrictive. We obtain the existence of a unique critical point of f which is characterized by

$$f(v_0) = \max_{x \in X} \min_{y \in Y} f(x + y) = \min_{y \in Y} \max_{x \in X} f(x + y).$$

* This work was sponsored by the Consejo de Desarrollo Cientifico y Humanistico de la Universidad de Los Andes.

† Present address: Gabriel D'Annunzio 7018, La Reina, Santiago, Chile.

0022-247X/82/110064–08$02.00/0

In [2], using a method which differs from ours, a certain generalization of Theorem 1 was given. Their results give conditions under which the existence of a unique critical point of f can be ascertained for the case when X and Y are not necessarily finite dimensional. Their conditions on the Hessian of f, however, are not as weak as those given here. Furthermore in [2] this critical point is not characterized as a min–max point of f.

2. The Main Theorem

We shall present our main theorem. First, we introduce some notation. If H and F are two real Hilbert spaces and $T: H \to F$ is a differentiable mapping, then $T'(x)$ denotes the derivative of T at $x \in H$. If $f: H \to \mathbb{R}$ is C^2, then $\nabla f(x)$ and $D^2 f(x)$ denote respectively the gradient and the Hessian of f at $x \in H$. In this case $\nabla f: H \to H$ is a C^1 mapping and $(\nabla f)'(x) = D^2 f(x)$ is a bounded self-adjoint linear operator on H.

The next two propositions are used in this work. For a proof of Proposition 1, see [6, p. 18]. The proof of Proposition 2 follows from Proposition 1 and from [5, Theorem 2.8].

Proposition 1. *Let H be a real Hilbert space and let L be a bounded linear operator on H. Suppose that*

$$\langle Lx, x\rangle \geqslant a\,\|x\|^2 \tag{1}$$

for all $x \in H$, where a is a positive real number. Then L is an isomorphism onto H and $\|L^{-1}\| \leqslant a^{-1}$.

Proposition 2. *Let H be a real Hilbert space and $T: H \to H$ a C^1 mapping. Let $\gamma: [0, \infty) \to (0, \infty)$ be a continuous nonincreasing function such that*

$$\int_1^\infty \gamma(s)\,ds = \infty, \tag{2}$$

$$\langle T'(y)k, k\rangle \leqslant \gamma(\|y\|)\,\|k\|^2 \tag{3}$$

for all $y \in H$ and all $k \in H$. Then T is a C^1 diffeomorphism of H onto itself.

Next we shall present our main theorem.

Theorem 2. *Let H be a real Hilbert space and let $f: H \to \mathbb{R}$ be of class C^2. Suppose that there exist two closed subspaces X and Y such that*

$H = X \oplus Y$ and two continuous nonincreasing functions α: $[0, \infty) \to (0, \infty)$, β: $[0, \infty) \to (0, \infty)$, such that

$$\int_1^\infty \alpha(s)\, ds = \infty, \qquad \int_1^\infty \beta(s)\, ds = \infty, \tag{4}$$

$$\langle D^2 f(x+y)k, k\rangle \geqslant \alpha(\|y\|)\,\|k\|^2 \tag{5}$$

for all $x \in X$, $y \in Y$, and $k \in Y$; and

$$\langle D^2 f(x+y)k, k\rangle \leqslant -\beta(\|x\|)\,\|k\|^2 \tag{6}$$

for all $x \in X$, $y \in Y$, and $k \in X$. Then

(a) *there exists a unique $v_0 \in H$, such that $\nabla f(v_0) = 0$;*

(b) $f(v_0) = \max_{x\in X} \min_{y \in Y} f(x+y) = \min_{y\in Y} \max_{x\in X} f(x+y)$.

Proof. (a) For each $x \in X$, let $g_x : Y \to \mathbb{R}$ be defined by $g_x(y) = f(x+y)$. Then $g_x \in C^2(Y, \mathbb{R})$, and for $k \in Y$

$$\langle \nabla g_x(y), k\rangle = \langle \nabla f(x+y), k\rangle \tag{7}$$

and hence in view of (5),

$$\langle D^2 g_x(y)k, k\rangle = \langle D^2 f(x+y)k, k\rangle \geqslant \alpha(\|y\|)\,\|k\|^2. \tag{8}$$

Then from Proposition 2 it follows that ∇g_x is a C^1 diffeomorphism from Y onto Y. Thus for each $x \in X$ there exists a unique $y_x \in Y$, such that $\nabla g_x(y_x) = 0$. Let us define $\phi: X \to Y$ by $y_x = \phi(x)$; then $\nabla g_x(\phi(x)) = 0$. We claim that $\phi \in C^1(X, Y)$. To see this, let P denote the orthogonal projection of H onto Y. Then it is easy to see that

$$y = \phi(x) \qquad \text{if and only if} \quad P\,\nabla f(x+y) = 0. \tag{9}$$

Next let us define $E: X \times Y \to Y$ by

$$E(x, y) = P\,\nabla f(x+y). \tag{10}$$

Then E is of class C^1 and given any pair $x_0 \in X$, $y_0 \in Y$ such that $E(x_0, y_0) = 0$, it follows that $y_0 = \phi(x_0)$. If E_y denotes the partial derivative of E with respect to y and if $k \in Y$, we have

$$E_y(x_0, y_0)k = PD^2 f(x_0 + y_0)k. \tag{11}$$

Clearly, $E_y(x_0, y_0)$: $Y \to Y$ and is a bounded linear mapping and we have from (11) and (5)

$$\langle E_y(x_0, y_0)k, k\rangle = \langle D^2 f(x_0 + y_0)k, k\rangle \geqslant \alpha(\|y_0\|)\,\|k\|^2, \tag{12}$$

for all $k \in Y$. From Proposition 1, $E_y(x_0, y_0)$ is an isomorphism onto Y. Then from the implicit function theorem [3, p. 115], there exists a C^1 mapping $\tilde{\phi}$ from a neighborhood U of x_0 in X into Y, such that, $E(x, \tilde{\phi}(x)) = 0$, for all $x \in U$. Moreover, from (9) and (10), $\tilde{\phi}(x) = \phi(x)$, for all $x \in U$. Hence since x_0 was arbitrarily chosen, it follows that $\tilde{\phi}$ can be defined over all of X. This enables us to conclude that $\phi \in C^1(X, Y)$.

Next let us define $G: X \to \mathbb{R}$ by

$$G(x) = f(x + \phi(x)). \tag{13}$$

Using the fact that ϕ is C^1, it follows that G is also C^1. However, we shall, in fact, show that G is a C^2 functional from X into $\mathbb{R}$. To accomplish this, note that for $h \in X$

$$\langle \nabla G(x), h \rangle = \langle \nabla f(x + \phi(x)), h + \phi'(x)h \rangle. \tag{14}$$

From (9)

$$\langle \nabla f(x + \phi(x)), k \rangle = 0 \tag{15}$$

for all $k \in Y$. Since $\phi'(x)h \in Y$, it then follows that

$$\langle \nabla G(x), h \rangle = \langle \nabla f(x + \phi(x)), h \rangle. \tag{16}$$

Let S denote the orthogonal projection of H onto X. Then S is bounded linear self-adjoint operator on H. For $h \in X$, $Sh = h$. Then from (16)

$$\langle \nabla G(x), h \rangle = \langle S \nabla f(x + \phi(x)), h \rangle, \tag{17}$$

for all $h \in X$. Equivalently,

$$\nabla G(x) = S \nabla f(x + \phi(x)). \tag{18}$$

From here we have that $\nabla G: X \to X$ is a C^1 mapping and therefore $G: X \to \mathbb{R}$ is a C^2 functional.

For $h \in X$,

$$\langle D^2 G(x)h, h \rangle = \langle D^2 f(x + \phi(x))h, h \rangle + \langle D^2 f(x + \phi(x))\, \phi'(x)h, h \rangle. \tag{19}$$

Recalling (15) we obtain,

$$\langle D^2 f(x + \phi(x))(h + \phi'(x)h), k \rangle = 0. \tag{20}$$

for all $x \in X$ and $k \in Y$. For the special case when $k = \phi'(x)h$, we can use the fact that for each $u \in H$ $D^2 f(u)$ is a self-adjoint operator on H, to obtain

$$\langle D^2 f(x + \phi(x))\, \phi'(x)h, h \rangle = -\langle D^2 f(x + \phi(x))\, \phi'(x)h, \phi'(x)h \rangle, \tag{21}$$

for all $x \in X$ and $h \in X$. Substituting (21) into (19) yields

$$\langle D^2G(x)h, h\rangle = \langle D^2f(x+\phi(x))h, h\rangle - \langle D^2f(x+\phi(x))\,\phi'(x)h, \phi'(x)h\rangle. \quad (22)$$

Then from (5) and (6),

$$\langle -D^2G(x)h, h\rangle \geqslant \beta(\|x\|)\,\|h\|^2 + \alpha(\|\phi(x)\|)\,\|\phi'(x)h\|^2 \geqslant \beta(\|x\|)\,\|h\|^2, \quad (23)$$

for all $x \in X$ and $h \in X$. Therefore by Proposition 2 $-\nabla G$ is a C^1 diffeomorphism from X onto X. Thus there exists a unique $x_0 \in X$, such that $\nabla G(x_0) = 0$. Then from (16) there exists a unique $v_0 = x_0 + \phi(x_0)$ such that

$$\langle \nabla f(v_0), h\rangle = 0 \quad (24)$$

for all $h \in X$. In light of (15),

$$\langle \nabla f(v_0), h + k\rangle = 0, \quad (25)$$

for all $h \in X$ and for all $k \in Y$. It follows that there exists a unique $v_0 \in H$, such that $\nabla f(v_0) = 0$.

(b) Let $k \in Y$ and $x \in X$, then by (15) and Taylor's formula [1, p. 211], there exists some $t \in (0, 1)$ such that

$$f(x+\phi(x)+k) - f(x+\phi(x)) = \tfrac{1}{2}\langle D^2f(x+\phi(x)+tk)k, k\rangle. \quad (26)$$

Then, from (5)

$$f(x+\phi(x)+k) - f(x+\phi(x)) \geqslant \tfrac{1}{2}\alpha(\|\phi(x)+tk\|)\,\|k\|^2. \quad (27)$$

Thus

$$f(x+\phi(x)) \leqslant f(x+\phi(x)+k) \quad (28)$$

for all $k \in Y$. Now given any $k \in Y$, we notice that one can always write

$$k = -\phi(x) + y \quad (29)$$

for some appropriate $y \in Y$. This implies that

$$f(x+\phi(x)) = \min_{y\in Y} f(x+y). \quad (30)$$

Again from Taylor's formula and using the fact that $\nabla G(x_0) = 0$, we have that for $h \in X$

$$G(x_0+h) - G(x_0) = \tfrac{1}{2}\langle D^2G(x_0+th)h, h\rangle \quad (31)$$

for some $t \in (0, 1)$. Using (23),

$$G(x_0 + h) - G(x_0) \leqslant -\tfrac{1}{2}\beta(\|x_0 + th\|)\,\|h\|^2, \tag{32}$$

for all $h \in X$. Then

$$G(x_0) = \max_{x \in X} G(x) = \max_{x \in X} f(x + \phi(x)). \tag{33}$$

Thus

$$f(v_0) = G(x_0) = \max_{x \in X} \min_{y \in Y} f(x + y). \tag{34}$$

From (5) and (6),

$$\langle -D^2 f(x + y)h, h\rangle \geqslant \beta(\|x\|)\,\|h\|^2 \tag{35}$$

and

$$\langle -D^2 f(x + y)k, k\rangle \leqslant -\alpha(\|y\|)\,\|k\|^2 \tag{36}$$

for all $x \in X$, $y \in Y$, $h \in X$, $k \in Y$. Repeating the same argument which was used to construct v_0, it follows that, there exists a unique $u_0 \in H$, such that, $-\nabla f(u_0) = 0$ and

$$-f(u_0) = \max_{y \in Y} \min_{x \in X} -f(x + y). \tag{37}$$

Then we must have that $v_0 = u_0$. Moreover, from (37),

$$f(v_0) = -\max_{y \in Y} \min_{x \in X} -f(x + y) = \min_{y \in Y} \max_{x \in X} f(x + y). \tag{38}$$

Thus, from (34) and (38)

$$f(v_0) = \max_{x \in X} \min_{y \in Y} f(x + y) = \min_{y \in Y} \max_{x \in X} f(x + y). \tag{39}$$

Remark. If in Theorem 1 we set $\alpha(s) = m_2$ and $\beta(s) = m_1$, $s \in [0, \infty)$, then it is easy to check that the conditions of Theorem 1 (even excluding X finite dimensional) imply the conditions of Theorem 2. Thus Theorem 1 follows from Theorem 2. Finite dimensionality of X is thus not needed by us.

EXAMPLE. Let X and Y be two closed subspaces of a real Hilbert space H, such that, $H = X \oplus Y$. Let T be a bounded linear self-adjoint operator on H and $V\colon H \to \mathbb{R}$ a C^2 functional. Suppose there exist constants μ_1, μ_2, ν_1, ν_2

such that $\mu_i > 0$, $i = 1, 2$, $\nu_1 \leqslant \nu_2$, and if $r = \min\{\mu_1, \mu_2\}$, then $\nu_i \in (-r, r)$, $i = 1, 2$. Furthermore suppose that

$$\langle Tz, z\rangle \leqslant -\mu_1 \|z\|^2 \qquad \text{for all} \quad z \in X, \tag{40}$$

$$\langle Tz, z\rangle \geqslant \mu_2 \|z\|^2 \qquad \text{for all} \quad z \in Y, \tag{41}$$

and

$$\nu_1 I \leqslant D^2 V(u) \leqslant \nu_2 I \qquad \text{for all} \quad u \in H, \tag{42}$$

I is the identity operator on H. Under these conditions. the equation

$$Tz + \nabla V(z) = b, \tag{43}$$

has a unique solution for each $b \in H$. Furthermore, if we define $f: H \to \mathbb{R}$ by

$$f(z) = \tfrac{1}{2}\langle Tz, z\rangle + V(z) - \langle b, z\rangle, \tag{44}$$

then f is a C^2 functional and the solution c of (43) is characterized by

$$f(c) = \max_{x \in X} \min_{y \in Y} f(x+y) = \min_{y \in Y} \max_{x \in X} f(x+y). \tag{45}$$

In fact, we have that

$$\nabla f(z) = Tz - b + \nabla V(z), \tag{46}$$

$$D^2 f(z) = T + D^2 V(z). \tag{47}$$

Let $m_1 = \mu_1 - \nu_2$ and $m_2 = \mu_2 + \nu_1$, then m_1 and m_2 are positive constants. From (40), (42), and (47) we obtain that for all $z \in H$ and $h \in X$

$$\langle D^2 f(z)h, h\rangle \leqslant -m_1 \|h\|^2 \tag{48}$$

and from (41), (42), and (47) we obtain that for $z \in H$ and $k \in Y$

$$\langle D^2 f(z)k, k\rangle \geqslant m_2 \|k\|^2. \tag{49}$$

Thus if we set $\alpha(s) = m_2$, $\beta(s) = m_1$, $s \in [0, \infty)$, we have that all the conditions of Theorem 2 are satisfied. It follows from this theorem the existence of a unique $c \in H$, such that

$$\nabla f(c) = Tc + \nabla V(c) - b = 0,$$

and such that (45) is satisfied.

ACKNOWLEDGMENT

The author is grateful to the reviewer for his comments which resulted in an improvement of this paper.

REFERENCES

1. R. G. BARTLE, "The elements of Real Analysis," Wiley, New York/London/Sydney, 1964.
2. P. W. BATES AND I. EKELAND, A saddle point theorem *in* "Differential Equations" (S. Ahmad, M. Keener, and A. C. Lazer, Eds.), Academic Press, New York, 1980.
3. M. S. BERGER, "Nonlinearity and Functional Analysis," Academic Press, New York/London, 1977.
4. A. C. LAZER, E. M. LANDESMAN AND D. R. MEYER, On saddle point problems in the calculus of variations, the Ritz algorithm and monotone convergence, *J. Math Anal. Appl.* **52** (1975), 594–614.
5. M. RADULESCU AND S. RADULESCU, Global inversion theorems and applications to differential equations, *Nonlinear Anal. Theory, Meth. Appl.* **4** (1980), 951–965.
6. J. T. SCHWARTZ, "Nonlinear Functional Analysis," Gordon & Breach, New York, 1969.